

DOCUMENT RESUME

ED 243 686

SE 044 417

AUTHOR Bull, Elizabeth Kay
TITLE Using Verbal Protocol Data to Reflect the Quality of Problem Representation in Solving Algebra Word Problems.
PUB DATE [84]
NOTE 3lp.
PUB TYPE Reports - Research/Technical (143)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Algebra; *Cognitive Processes; *College Mathematics; Educational Research; Error Patterns; Higher Education; *Mathematics Instruction; Problem Sets; *Problem Solving; *Protocol Materials
IDENTIFIERS *Mathematics Education Research

ABSTRACT

The goal of this study was to find a way to quantify three criteria of representational quality, described by Greeno, so that it would be possible to examine statistically the relationship between representational quality and other variables related to problem solution. The sample consisted of 18 college students, 84 percent of whom had mathematics through trigonometry or calculus in high school as well as at least one college mathematics course. Students were asked individually to think aloud as they solved six algebra word problems, and then to give a retrospective report describing what could be remembered about the problem-solving process used. Each idea unit in the protocols was coded, using a slightly modified coding system developed by Ericsson and Simon. Then, coherence and correspondence scores were computed for each problem. Both scores were found to predict whether or not a student would be able to solve correctly at least some word problems, with the correspondence score a stronger predictor. Three types of errors made by students were also noted: interpretive, integrative, and recall errors. Appendices contain the word problem set, a sample coded verbal protocol, and a task analysis for one problem. (MNS)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

ED243686

Using Verbal Protocol Data to Reflect the Quality of Problem
Representation in Solving Algebra Word Problems

Elizabeth Kay Bull

Moorhead State University

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

✓ This document has been reproduced as
received from the person or organization
originating it.
☐ Minor changes have been made to improve
reproduction quality.

• Points of view or opinions stated in this docu-
ment do not necessarily represent official NIE
position or policy.

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

Elizabeth Kay Bull

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

Running Head: Verbal Protocol Data

A portion of the research reported in this paper was
completed as part of the author's doctoral dissertation.

Peter Polson and Anders Ericsson gave valuable suggestions
and support on this project. Lee Willer, Paula Lowe, Ken
Schneider, and Deb Anderson were most helpful in completing the
tedious task of coding the verbal protocols.

SE044417

Abstract

Based on the theory that success in solving mathematical word problems depends upon the quality of the underlying representation of the problem, verbal reports were used to develop a means for quantifying the quality of the representation. Measures of the coherence of the representation and the correspondence of the representation to the demands of the task and the subject's store of mathematical and real world knowledge were correlated with whether or not subjects correctly solved each of six algebra word problems. Correspondence emerged as a stronger predictor of success than did coherence for most of the problems.

Using Verbal Protocol Data to Reflect the Quality of Problem
Representation in Solving Algebra Word Problems

Elizabeth Kay Bull

Moorhead State University

Both mathematics educators and cognitive psychologists have become increasingly concerned over the past decade with detailed examinations of the problem solving processes used in attacking mathematical word problems. One way such processes have been examined is through the use of verbal (or think aloud) protocols. A verbal protocol is produced by asking subjects to think aloud while solving a given problem and tape recording the resulting verbalizations. These are transcribed and analyzed. The method of analyzing the protocol varies with the researcher's field of study.

Researchers in mathematics education have typically analyzed verbal protocols by using the coding system developed by Kilpatrick (1967). Kilpatrick's system is based on Polya's (1957, 1962) descriptions of heuristic strategies and is used to categorize which of several heuristics (e.g., "guessing," "using an equation," "algorithms") were being used by subjects in solving problems, to determine how often each heuristic is used, and to correlate heuristics use with success in solving the problem.

On the other hand, while cognitive psychologists have made some attempts at systematically coding verbal protocols (Ericsson & Simon, in press), they have more often used the protocols to develop detailed models or computer simulations of the subject's solution process (e.g., Greeno, 1976; Newell & Simon, 1972). One

aspect of the problem solving process which has been identified through these procedures is that of "problem representation." The problem representation is the way in which a subject thinks about or organizes a problem so that possible solution methods can be considered. Previous research on problem representation using physics problems (e.g., Larkin, 1979) and mathematics problems (e.g., Greeno, 1978; Krutetskii, 1976; Resnick & Ford, 1981) has indicated that the process of building a representation is elaborative. That is, the problem solver begins with bits of information gained through understanding the problem text and integrates these with inferences about the problem situation and additional information recalled from memory. Integration and elaboration continue until the representation is complete enough to allow for solution of the problem. Thus, while a representation is formed at the beginning of the solution attempt, it is repeatedly modified throughout the process of solving the problem.

Attempts at improving the problem-solving skills of mathematics students have frequently centered on teaching students to use the heuristic strategies described by Polya (1957). The results of such efforts have met with only limited success (e.g., Bassler, Beers, & Richardson, 1975). Interpreting such results from an information-processing viewpoint, one could say that this outcome is not surprising because a heuristic strategy would be useless without an underlying integrated representation of the problem. In addition, simply identifying the series of heuristics used to solve a problem provides limited information about the

characteristics of successful, as compared with unsuccessful problem solving attempts. Thus, a preliminary step in the development of a coherent theory of mathematical word problem solving would need to go beyond looking at the heuristic strategies used by problem solvers as done by Kilpatrick (1967) and other mathematics educators and examining in detail the quality of the underlying problem representation.

Since the problem representation is a cognitive structure and is thus essentially within a person's mind, two questions arise for researchers in this field: (1) How can we know that a representation exists? and (2) How can we measure or evaluate it? The answer to the first requires the simple leap of faith of a cognitive psychologist; that is, we know that it exists. Ericsson and Simon (in press) argue that the problem solver cannot act on a problem unless the actions necessary are available to the subject within the working memory store. It is the contents of this working memory store which constitute the representation. In answer to the second question, we must rely on what the problem solver tells us about the solution process and infer the structure of the representation from this information, since information not in working memory cannot be verbalized (Ericsson & Simon, in press).

Assuming then that problems are represented cognitively by the problem solver, what do we know about the quality of the problem representation? Our information comes from two types of psychological research--developmental studies and studies comparing the solution paths used by experts and with those used

by novices. The developmental literature in problem solving often uses Piagetian-type tasks and examines the changes in the strategies used by children of different ages. For example, Siegler (1976; Klahr & Siegler, 1978) identified a series of four performance levels for children aged five to nine on a balance scale task and concluded that the growing ability to handle the task with increasing age could be traced to increasingly richer and more accurate encodings (representations) of the problem situations.

Larkin and her colleagues (Larkin, 1979; McDermott, Simon & Simon, 1980) have examined differences in problem representations between experts and novices in solving physics problems. The conclusions reached by these researchers indicate that (a) experts have more knowledge than do novices, (b) the expert's knowledge is stored as chunks such that a whole schemata is often recalled in solving a given type of problem, and (c) experts use physical diagrams and therefore produce more integrated representations of the problem than do the novices. Clement (1979) further suggests that in solving physics problems the novice engages in more search for a solution path and uses direct translation strategies which lead to one-to-one mapping between the problem text and the solution procedures.

How then can we measure the quality of the underlying problem representation so that we know whether or not there are differences in the representations formed by novices and experts in solving the problem? According to Greeno (1977), the quality of a problem representation can be described in three ways:

1. How coherent, or well-integrated, the various parts of the representation are with each other. A poor representation is based on loose connections between the parts while a good representation contains parts which are connected to each other.

2. How well the representation corresponds to the actual problem text. The more accurately the representation reflects the demands of the task, the more likely the subject is to solve the problem.

3. The degree to which the representation is connected to the subject's knowledge store. The knowledge store contains specific knowledge relevant to the problem being solved, procedural knowledge regarding problem solving procedures, and general world knowledge. The more relevant and correct the information that is recalled, the better the representation is. Without this knowledge, the representation is incomplete and/or superficial.

Greeno (1977) operationalized these criteria of good problem representation--or understanding--through the use of a computer model of geometry problem solving. He did not, however, attempt to operationalize them so that they are quantifiable. Thus, it was the goal of the current study to find a way to quantify the three criteria of representational quality as described by Greeno--coherence, correspondence, and connectedness--so that it would be possible to statistically examine the relationship between representational quality and other variables related to problem solution.

Method

Subjects

The sample included ten males and eight females drawn from the general psychology subject pool at the University of Colorado. Approximately half of the subjects were majoring in engineering, science, or mathematics, and an additional 28% were business majors. They appeared to be a mathematically capable group with 84% of the subjects having had strong high school training in mathematics (through trigonometry or calculus) as well as at least one college level mathematics course.

Materials

Six algebra word problems, drawn from the research literature and similar to those contained in a high school algebra textbook, were used (see Appendix A). One of the problems was impossible to solve (from Paige & Simon, 1966) and was included to see whether subjects directly translated the problem into a mathematical equation or could examine a problem logically and thus spot the incongruity. Because of the special nature of this problem, it was not included in the analyses.

Procedure

Subjects were seen individually by the experimenter. Each was asked to complete a general information questionnaire and was then instructed in thinking aloud procedures. After solving two practice problems, each subject was presented with the six algebra word problems, one at a time, and asked to solve each. After each problem was solved, the subject was asked to give a retrospective report describing what could be remembered about the problem

solving process used. No time limits were imposed, and the entire process took between one and two hours.

Scoring the Verbal Protocols

The transcribed protocols were divided into idea units in a manner similar to the procedure described by Caccamise (1981) in which the natural language structures of the protocols are maintained. Idea units--or propositions--as used here, are short statements corresponding to separate, yet complete, ideas. (See Appendix B for an example of a protocol which has been broken into idea units.)

Few generic verbal protocol coding systems exist in the literature. One which seemed to be applicable to the data obtained in this research and to the purposes of this project was developed by Ericsson and Simon (in press). This system was designed to distinguish among information presented in the problem (literal statements), information held in working memory (clarifications), and information generated by the subject in response to the problem-solving activity (generations). Slight modifications of the Ericsson and Simon system were made for this research such that each idea unit in the protocol was coded as belonging to one of the following categories:

1. Literal Statements (L)--instances in which the exact wording of the problem text was repeated; i.e., the subject was reading or rereading the problem.

2. Clarifications (C)--attempts by the subject to clarify or understand the demands of the problem text by rephrasing the

information given, drawing inferences about the nature of the problem, writing down or underlining information in the text, etc.

3. Planning Statements (P)--statements or short pauses indicating that planning or search was taking place or goals were being set.

4. Solution Generations (S)--statements related to the solution procedure itself including introducing variables, establishing relationships, writing and solving equations, performing calculations, etc.

5. Evaluative Statements (E)--statements indicating that an inference, procedure, or answer was being evaluated.

6. Other Verbalizations (O)--miscellaneous verbalizations or pauses that did not fit within any of the other categories.

A copy of the coding manual is contained in Bull (1982) or can be obtained from the author. The protocol in Appendix B demonstrates the use of this coding system.

Inter-rater and intra-rater reliabilities were established using randomly selected protocols. Agreements on individual protocols ranged from 58% to 93% with agreement lowest on those protocols containing large numbers of statements coded as "clarifications", "planning statements", and "other verbalizations." Average agreement across all of the protocols was 80%.

Calculation of Coherence, Correspondence, and Connectedness Scores

Coherence. According to Greeno (1977), a coherent representation has clearly established goals and develops a

high-level strategic plan for driving the solution. A coherent representation contains parts which are well-integrated, requires little search for a solution path, and contains few repetitions of ideas or circular passes through the problem space. Based on these assumptions, a coherence score was computed for each problem by combining two measures of integration: (a) the proportion of the total number of idea units in the protocol coded as "solution generations," and (b) the proportion of the total amount of time the subject worked on the problem during which the subject did not verbalize, i.e., pauses. The coherence score was then calculated by subtracting the resulting time score (b, above) from the proportion of solution generations score (a, above). Thus, the greater the difference between these two proportions, the higher the coherence score, or the more time spent in actually solving the problem and the more goal directed the solution process seemed to be. The highest possible score was 1.00. Given the subtraction effect, it was possible for protocols to receive a negative score--especially if there were excessively long pauses in the protocol. Figure 1 contains two samples of how the coherence scores were calculated. WD, who obtained the higher score, spent considerably more time than did YP in the actual process of solving the problem. YP, on the other hand, spent more time searching for a workable solution plan and evaluating the various parts of her plan. (YP's full protocol is provided in Appendix B.)

Subject	YP		WD	
Idea Units	Number	Proportion (Number/Total) in Protocol	Number	Proportion (Number/Total) in Protocol
Literal Statements	12	.185	7	.156
Clarifications	10	.154	1	.022
Planning Generations	10	.154	1	.022
Solution Generations	22	.338	32	.711
Evaluative Generations	5	.077	3	.067
Other	6	.092	1	.022
TOTAL IDEA UNITS	65	--	45	--
Subscore A = (Proportion Solution Generations)	--	A = .338	--	A = .711
	Time in Solution	Pauses	Time in Solution	Pauses
Time Measures	236"	47"	135"	15"
Subscore B = Time in pauses/ total time in solution attempt)	--	B = .199	--	B = .111
COHERENCE SCORE (C) C = A - B	--	C = .139	--	C = .600

Figure 1

Sample Calculation of Coherence Scores
Dimes/Nickels Problem

Correspondence and Connectedness. Greeno (1977) defines correspondence as the degree to which the representation corresponds to the actual problem text and connectedness as the degree to which the representation is connected to the problem solver's general knowledge store. While these two concepts are easily distinguishable in theory, operationally they were very difficult to separate. Therefore, the correspondence score was computed to include both of these concepts. First, a task analysis was performed to include all of the steps necessary for the correct solution of each problem. For the dimes/nickels problem, for example (see Appendix C), there are two correct paths that subjects may take--one reflects an equation-writing path and the other a non-algebraic strategy of generating and testing values for the coins. In addition, a set of "critical flags" was identified for each problem. A critical flag was defined as real world knowledge (e.g., that there cannot be a negative or fractional number of coins) that should have been heeded (see Appendix C). The correspondence score for each problem reflected (a) the number of items in the task analysis that were included in the solution (b) minus the number of critical flags that should have been heeded but weren't (c) divided by the total number of possible items in the task analysis.

For example, sample subject YP used only ten of the 23 critical propositions in the dimes/nickels problem and raised no critical flags; she therefore received a correspondence score of 10/23 or .435. A high correspondence score reflected a representation having a high degree of correspondence to the

actual problem text and to correct mathematical procedures plus a high degree of connectedness to the subject's general knowledge store.

Results

Each subject received several scores for each problem: (a) whether or not the correct answer was obtained, (b) whether or not a correct equation was written for the problem, (c) the time needed to solve the problem, (d) proportions for each of the categories of idea units, (e) a correspondence score, and (f) a coherence score. Table 1 provides summary information for several of the variables above for each of the six problems in the set.

Problem Solution. The easiest problems in the set were (a) the dimes/nickels problem which was correctly solved by 83% of this sample and (b) the pool problem, solved by 78% of the subjects. The jet problem (solved by only 22% of the subjects) and the impossible coin problem (solved by 33%) were the most difficult problems in the set.

While all of the problems--except the impossible coin problem--could be solved algebraically, it was possible to solve the dimes/nickels, pool, and age problems using non-algebraic strategies. The coin and age problems were often solved by generating trial answers and testing these within the constraints of the problem text. This strategy was more effective in solving some of the problems than others. For example, subjects who easily solved the mother/daughter age problem through trial and error sometimes had difficulty with the mother/son age problem.

Table 1

Summary Statistics for Solutions to Algebra Word Problems

Problem	Percent solving problem correctly	Percent writing correct equation	Mean time to solve problem	Mean coherence score	Mean correspondence score
Pool	78	6	4'32"	.12	.86
Dimes/ Nickels	83	61	5'28"	.38	.79
Mother/ Daughter	67	11	5'08"	.15	.57
Mother/ Son	56	33	5'27"	.22	.64
Jet	22	17	7'0"	.11	.45
Impossible Coin	33	--	9'32"	--	--

The majority of the subjects who were successful in solving the pool problem used non-algebraic methods; that is, they performed a series of arithmetic operations on the data presented in the problem text. There seems to be no need to write an equation for this problem since the quantity of information provided in the problem text is sufficient for number (rather than symbol) manipulation. If algebra is seen as a means of reducing complexity in a word problem, it is unnecessary where there is little complexity. The opposite situation is true for the jet problem which is quite complex. Algebraic reduction of that complexity in the form of an equation is essential for correct solution.

Writing an equation was only a partially successful strategy for the impossible coin problem. The problem was considered to be solved only if the subject noted the incongruity in the problem text. A few of the subjects did this immediately, while the others who noted the incongruity did so only after writing and solving an equation and coming up with an impossible solution--a negative or fractional number of coins.

Measures of Representational Quality. Mean coherence scores ranged from a low of .11 on the jet problem to a high of .38 for the dimes/nickels problem. Recall that this score was computed as a function of the amount of time actually spent in trying to solve the problem. For example, approximately half of the idea units in the average dimes/nickels protocol were coded as solution generations, while only 27% of the idea units in the jet protocols were so coded. The protocols suggest that subjects spent a considerably greater proportion of their time reading and trying to understand the jet problem and then searching for a solution plan than they did in solving the dimes/nickels problem.

While the impact of comprehension and planning processes varied by problem and by representational measure, in general it can be concluded that the more time a subject spent engaged in reading the problem text, trying to comprehend it, and searching for a solution plan, the less coherent the representation and the less it corresponded to the demands of the task environment. This did not necessarily mean that subjects who engaged in these processes were unable to solve the problem, but they were less likely to be able to solve it than were those who could more

quickly form a well-integrated and complete problem representation.

Correspondence scores were highest for the pool problem ($\bar{X}$ = .86; range: .30 to 1.00) and the dimes/nickels problem ($\bar{X}$ = .79; range: .45 to .95) and lowest for the jet problem ($\bar{X}$ = .45; range: .09 to .96). Since the jet problem could only be solved through the use of an equation, subjects were required to make more difficult inferences (e.g., the distance is equal in both situations) and recall more mathematical knowledge (e.g., $d = rt$) in the solution of this problem than was required in solving the other problems.

Measures of coherence and correspondence were significantly and positively correlated with successful problem solution as measured by correct/incorrect answer. Correlations between correct answer and correspondence score ranged from .46 to .77 for the various problems (see Table 2). All of these correlations were significant at the .05 level or better. Correlations between correct answer and coherence score were considerably weaker for most of the problems and ranged from .12 to .56. Only two of these were significant at the .05 level. Finally, the correlations between the two measures of representational quality--coherence and correspondence--were strong (.77, $p \leq .001$) for only one problem--the jet problem which was the most mathematically complex problem in the set.

For the jet problem, the correspondence of the representation to the subject's general knowledge store was a better predictor of success than was coherence of the representation--possibly because

Table 2

Correlations between Coherence and Correspondence
Scores and Correct Solution

Problem	Coherence Score/ Correct Answer	Correspondence Score Correct Answer	Coherence/ Correspondence
Pool	.098	.773 ^d	.088
Dimes/Nickels	.562 ^b	.465 ^a	.543 ^b
Mother/Daughter	.188	.746 ^d	.258
Mother/Son	.221	.487 ^a	.263
Jet	.540 ^a	.541 ^a	.777 ^d

^ap < .10
^bp < .05
^cp < .01
^dp < .001

the only way the problem can be solved is through the use of an integrated equation which is, in effect, measured by the correspondence score. The opposite is true for the dimes/nickels problem--coherence of the representation is the best predictor of successful solution. Since this problem can be solved equally well by writing an equation or using non-algebraic strategies, correspondence to mathematical knowledge is not as critical for the solution of this problem as it is for the jet problem.

Conclusions

Can representational quality in mathematical problem solving be measured quantitatively? The answer is a qualified yes. Coherence and correspondence scores were calculated to reflect the

correctness of the inferences made, the correctness of the information recalled from memory, the correctness of the procedures used, and the degree of integration of the solution process. Both the correspondence and coherence scores, as operationalized here, predict whether or not the subject will be able to correctly solve at least some word problems. The correspondence score, which is basically a measure of correctness, incorporates measures of the correctness of the inferences made, the information recalled from memory, and the procedures used. In general, the correspondence score was a stronger predictor of solution success than was the coherence score.

The coherence score, which reflects the amount of time actually spent solving the problem, had much lower predictive validity since it did not take into account the correctness of the procedure used. At least two different situations influenced the relationship between coherence score, correspondence score, and correct answer. First, some subjects were able, after lengthy explorations for an appropriate solution path, to find a way to solve the problem; thus, their protocols received low coherence scores even though they eventually found the correct answer. In these cases the correspondence score was usually high. In the second situation, subjects spent a great deal of their time following an incorrect solution path and thus received low correspondence scores and failed to find the correct answer. They did, however, receive high coherence scores because they were engaged in a solution process--even though it was an incorrect one.

While the research reported here was exploratory in nature, it did demonstrate that it is possible to quantify measures of representational quality in a reliable and valid manner. Since the number of subjects and problems used in this study was quite limited, the procedures developed must be further evaluated before being used to build a theoretical base for understanding the role of problem representation in the problem-solving process.

At this point, I wish to say something about the use of verbal protocols as a method of collecting and analyzing data. There were several difficulties associated with research such as that described in this paper. While measures of inter-rater reliability were adequate, it often seemed inordinately difficult to agree on parcelling protocols into idea units and even more difficult to specify the criteria used in coding the idea units. Repeated discussion and considerable training were needed to produce reliable scoring. A second, and probably greater, problem with verbal protocols concerns time requirements. The transcribing and coding of verbal data is extremely time consuming. It is doubtful that such efforts result in powerful enough data to be worth such large investments of time. However, verbal protocols do contain a wealth of information about problem-solving processes and deserve to be examined in detail. For example, use of the task analysis and critical flag data could be very helpful for educators trying to determine where errors in the problem-solving process have been and are most likely to be made. Improving instruction in problem solving will be impossible

without a careful delineation of the errors students make in the solution of mathematical word problems (Resnick, 1981).

Three types of errors were apparent in these verbal protocols: interpretative errors, integrative errors, and recall errors. Interpretative errors were caused by one of two problems--subjects either did not comprehend what the problem constraints were or were unable to effectively map the problem text onto a correct equation. For example, using direct translational strategies, subjects would read the phrase (in the mother/daughter problem) "ten years from now she will be only twice as old" and would write " $M + 10 = 2d$." They often failed to account for the daughter also aging ten years.

Integrative errors were noted in all problems but were particularly pronounced in the jet problem. Several subjects were able to recall all the necessary background information and to write two correct equations, but were unable to integrate them in order to solve the problem.

The third type of difficulty the poor problem solvers had was their failure to understand basic mathematical concepts or to recall critical mathematical knowledge. For example, some subjects seemed not to understand the concept of equality and wrote equations like " $4x = x$ " to represent the relationship between the mother's and son's ages, or they failed to recall that variables stand for distinct entities, and used " x " to stand for both the number of nickels and the number of dimes (e.g., " $.10x + .05x = 1.60$ ") in the coin problem.

One final point: much of the research in teaching mathematical problem-solving skills has attempted to teach students to use general heuristic strategies. Since results of this body of work suggest that teaching such general strategies does not improve problem-solving skill, perhaps we need to turn our research attention to improving more specific skills; such as how to form an integrated problem representation, interpret mathematical concepts in light of a specific problem, and read and understand problem texts so that they can be translated into mathematical symbols in meaningful units rather than through the direct translation of key words into mathematical symbols. This project represents a beginning in this direction, but there is still much to be done in developing a coherent theory of problem solving in mathematics.

References

- Bassler, O. C., Beers, M. I., & Richardson, L. L. (1975). Comparison of two instructional strategies for teaching the solution of verbal problems. Journal for Research in Mathematics Education, 6, 170-177.
- Bull, E. K. (1982). Problem representations and solution procedures used in solving algebra word problems. Unpublished doctoral dissertation, University of Colorado.
- Caccamise, D. J. (1981). Cognitive processes in writing: Idea generation and integration. Unpublished dissertation, University of Colorado.
- Clement, J. (1979). Mapping a student's casual connections from a problem-solving protocol. In J. Lochhead & J. Clement (Eds.), Cognitive process instruction. Philadelphia: Franklin Institute Press.
- Ericsson, K. A., & Simon, H. A. (in press). Verbal reports as data.
- Greeno, J. G. (1976). Indefinite goals in well-structured problems. Psychological Review, 83, 479-491.
- Greeno, J. G. (1977). Processes of understanding in problem solving. In N. J. Castellan, Jr., D. B. Pisoni, & G. R. Potts (Eds.), Cognitive theory (Vol. 2). Hillsdale, N. J.: L. Erlbaum.
- Greeno, J. G. (1978). A study of problem solving. In R. Glaser (Ed.), Advances in instructional psychology (Vol. 1). Hillsdale, N. J.: L. Erlbaum.

- Kilpatrick, J. (1967). Analyzing the solution of word problems in mathematics: An exploratory study (Doctoral Dissertation, Stanford University, 1967). Dissertation Abstracts International, 28, 4380A.
- Klahr, D., & Siegler, R. S. (1978). The representation of children's knowledge. In H. W. Reese & L. P. Lipsitt (Eds.), Advances in child development and behavior (Vol. 12). New York: Academic Press.
- Krutetskii, V. A. (1976). The psychology of mathematical abilities in schoolchildren. (J. Teller, trans., J. Kilpatrick & I. Wirzup, eds.) Chicago: University of Chicago Press.
- Larkin, J. H. (1979). Information processing models and science instruction. In J. Lochhead & J. Clement (Eds.), Cognitive process instruction. Philadelphia: Franklin Institute Press.
- Larkin, J., McDermott, J., Simon, D. P., & Simon, H. A. (1980). Expert and novice performance in solving physics problems. Science, 208, 1335-1342.
- Newell, A., & Simon, H. A. (1972). Human problem solving. Englewood Cliffs, N. J.: Prentice Hall.
- Paige, J. M. & Simon, H. A. (1966). Cognitive processes in solving algebra word problems. In B. Kleinmuntz (Ed.), Problem solving: Research, method, and theory. New York: John Wiley & Sons.
- Polya, G. (1957). How to solve it (2nd Ed.). Garden City, N. Y.: Doubleday.

- Polya, G. (1962). Mathematical discovery. On understanding learning and teaching problem solving. (Vol. 1). New York: John Wiley & Sons.
- Resnick, L. B. (1981). Instructional psychology. Annual Review of Psychology, 32, 659-704.
- Resnick, L. B. & Ford, W. M. (1981). The psychology of mathematics or instruction. Hillsdale, N. J.: L. Erlbaum.
- Robinson, M. S. (no date). Solving algebra word problems: Knowledge and representation. Unpublished manuscript. Pittsburgh: Carnegie-Mellon.
- Siegler, R. S. (1976). Three aspects of cognitive development. Cognitive Psychology, 4, 481-520.

Appendix A

Word Problem Set

1. To be filled in 24 hours, a pool must be filled at the rate of 94 gallons per hour. The main water pipe, through which water can flow at the rate of 64 gallons per hour is turned on. However, the pool directors learn that this will not get the pool done in time. So, in order to supplement the first, they turn on a second, smaller pipe for a thirteen hour period. At what rate must water flow in the second hose if the pool is to be filled on time? (from Robinson, no date)
2. A boy has some dimes and nickels. There are 21 coins worth \$1.60. How many dimes and how many nickels does the boy have? (from Kilpatrick, 1967)
3. A mother is 3 times as old as her daughter; 10 years from now she will be only twice as old. What is the mother's age? (from Krutetskii, 1976)
4. A mother is four times as old as her son. In four years she will be three times as old as he is. How old is the mother now? (developed for this set)
5. An experimental jet is tested in a wind tunnel. Flying against the air stream it travels the length in 10 minutes. The next day it is tested flying with the air stream and covers the length of the wind tunnel in 6 minutes. If the speed of the wind in the tunnel is 15 miles per hour, what speed can the jet fly with the wind turned off? (from Robinson, no date)
6. The number of quarters a man has is seven times the number of dimes he has. The value the dimes exceeds the value of the quarters by two dollars and fifty cents. How many has he of each coin? (Paige & Simon, 1966)

Appendix B

Sample Coded Verbal Protocol

Note: The following symbols were used to represent subject non-verbal behavior.

- (w:) indicates that the subject wrote something
 (adds) indicates that something was added to an already started equation
 (c) indicates use of the calculator
 . . . each dot represents a one-second pause

PROBLEM: Dimes/Nickels

SUBJECT: YP (Female, science major)

<u>Idea Unit</u>	<u>Code</u>	<u>Verbalization</u>
1	L	A boy has some dimes
2	L	and nickels.
3	L	There are 21 coins
4	L	worth \$1.60.
5	L	How many dimes
6	L	and how many nickels
7	L	does the boy have?
8	C	. . 21.
9	L	How many dimes
10	L	and how many nickels
11	L	does the boy have?
12	C	Altogether he has 21.
13	O	(laughs)
14	O	OK, um, . . .
15	C	there's a total of 1.60
16	C	and divided into dimes and nickels
17	S	which means there's 2 variables . . .
18	C	21 coins altogether
19	C	Hum, 21 coins and .
20	L	how many dimes
21	C	and nickels
22	L	does the boy have?
23	S	. . He could have any number of either of them.
24	C	Oh, but he's gotta have 21 total. Ok.
25	C	. . . 21 . . .
26	E	but he can't have 21 dimes
27	S	because then he'd have \$2.10.
28	E	. . And, he can't have 21 nickels.
29	O	They're not gonna make it that easy. .
30	C	21 total . . so, . . .
31	P	Oh, let's do this the baby way.

32 S Take . . . 5 . .
33 P There's gotta be a logical way to do
this.
34 S . . 20,
35 S 15 dimes, \$1.50.
36 E Can't have that many . .
37 O Hum, ... (unintelligible)
38 S Ok, 10 dimes, you'd have a dollar.
39 S And so \$1.00 equals 10 d (w)
40 S So, then he'd need 60 cents
41 S which would be 11, 12 nickels (w: .60 =
12n)
42 S which added together would make it a
total of 22 dimes and nickels.
43 E But there isn't . .
44 P and I need one less coin.
45 S So, if I take one less dime
46 S I'd have 90 cents (w: .90 = 9d)
47 P But, then I'd need .
48 P I'd have an unaccounted 10 cents.
49 P So, that would mean, .
50 P that might work.
51 S 90 cents worth of dimes
52 S and then, um, 70 cents worth of nickels
(w: .70 =)
53 S 70 divided by 5 (c) equals 14, 14 nickels
(adds: 14n)
54 S Add 14 and 9, get 23.
55 E I went the wrong direction.
56 P So I want to add another dime
57 P since I came out with one more total than
I needed.
58 S So, if I add . . \$1.10 worth in dimes
59 S which means 11 dimes (w: 1.10 = 11d)
60 S and only 50 cents worth of nickels,
(w: .50 = 5n)
61 P Which would make it,
62 O hum, this makes sense.
63 S You have 5 nickels
64 S and 11 dimes (w and circles)
65 O That's what you've got.

Appendix C

Task Analysis - Measures of Correspondence Dimes/Nickels Problem

NOTE: The algebraic solution pathway is represented by the items in the "S" column. Those aspects of the non-algebraic solution pathway which are different from the algebraic pathway are represented by the items in the "S-prime (S')" column.

- | | | |
|-----|---|--|
| S01 | A dime is worth 10 cents | |
| S02 | A nickel is worth 5 cents | |
| S03 | There are two variables | |
| S04 | The number of dimes = d | S'04 Test value generated for d |
| S05 | The number of nickels = n | S'05 Test value generated for n |
| S06 | There must be two equations | S'06 There must be two test constraints. |
| S07 | One equation for number of coins | |
| S08 | $d + n = 21$ | |
| S09 | One for value of coins | |
| S10 | The value of dimes = $.10d$ | S'10 Test: $(d)(.10) = a$ |
| S11 | The value of nickels = $.05n$ | S'11 Test: $(n)(.05) = b$ |
| S12 | Thus, $.10d + .05n = 1.60$ | S'12 $\$1.60 - a = b$
(or) $\$1.60 - b = a$ |
| S13 | Solve for one variable in terms of the other, e.g., $n = 21 - d$ | S'13 $21 - n = d$ (or) $21 - d = n$ |
| S14 | Substitute into other equation, e.g., $.10d + .05(21 - d) = 1.60$ | S'14 $b/.05 = n$ (or) $a/.10 = d$ |
| S15 | Multiply terms in () | |
| S16 | Combine all variable terms | |
| S17 | Combine all numeric terms | |
| S18 | Divide terms by constant | |
| S19 | Find d (or n) | S'19 Does $a + b = \$1.60$? |

S20 Substitute to find n (or d) S'20 Does $d + n = 21$?

S21 Answer, $n = 10$

S22 Answer, $d = 11$

CRITICAL FLAGS

1. Decimals have to be consistent across the money terms.
2. The value of the coins must be positive.
3. The number of coins must be a whole number.